



Das Vokabular zwischen Sprach- und Informationswissenschaft

Engerer, Volkmar Paul

Published in:
Sprachen verbinden

Publication date:
2018

Citation for published version (APA):
Engerer, V. P. (2018). Das Vokabular zwischen Sprach- und Informationswissenschaft. In V. Janikova, A. Brychova, J. Velickova, & R. Wagner (Eds.), *Sprachen verbinden: Beiträge der 24. Linguistik- und Literaturtage, Brno/Tschechien, 2016* (pp. 433-444). Verlag Dr. Kova. Sprache und Sprachen in Forschung und Anwendung (SiFA) Vol. 6

Das Vokabular zwischen Sprach- und Informationswissenschaft

Volkmar Engerer, Universität Kopenhagen, Aalborg, Dänemark

The vocabulary between linguistics and information science

This paper explores the relationships between natural language lexicons in lexical semantics and thesauri in information retrieval research. These different areas of knowledge have different restrictions on use of vocabulary. The differences in vocabulary properties correlate with two other factors, the information scientific dimension of Control (deliberate, social activities of building and maintaining vocabularies), and Syntagmatization, which is a well-known category from linguistics and language theory. It is proposed that there is an inverse relationship between Control and Syntagmatization.

Keywords: information science, linguistics, vocabulary, lexicon, lexical semantics, thesaurus, information retrieval, control, syntagmatization, diachrony, synchrony, languages for special purposes, terminology research

1 Sprach- und Informationswissenschaft, diachron und synchron – eine kleine Publikationsgeschichte

Volkmar Engerer, 56, ausgebildeter Linguist und seit 2012 in der Informationswissenschaft tätig. Ich stehe deshalb auch zwischen zwei Lagern, der Sprach- und der Informationswissenschaft. Diese beiden Disziplinen haben eine ganze Reihe von terminologischen Ausdrücken gemeinsam, z. B. *Bedeutung*, *Information*, *Text* und, nicht zuletzt, *Vokabular*, das Thema meines Beitrags. Dass ein gemeinsamer Wortschatz nicht immer auch eine Gemeinsamkeit der Referenten garantiert, ist nichts Besonderes, auch nicht in den Wissenschaftssprachen. Ich schätze deshalb die Augenblicke, wenn ich mit meinem Kollegen aus der Informationswissenschaft über das Gleiche oder etwas Vergleichbares spreche. Dann können wir auch ohne uns gleich vor den Kopf zu stoßen die Unterschiede beim Namen nennen. Das fördert den interdisziplinären Dialog.

Wie stehen linguistische und informationswissenschaftliche Perspektiven zueinander, wenn sie Bedeutung, Information, Text oder Vokabular ins Blickfeld nehmen? Um hier klüger zu werden, begann ich mit einem Rückblick, beginnend mit der Zeit ca. seit 1950, in der sich die Informationswissenschaft gerade als technologisch infizierte, moderne Naturwissenschaft zu etablieren suchte (Tredinnick 2006). Mein Ergebnis: Angewandte Linguisten hatten den computergetriebenen Informationswissenschaftlern viel zu geben, aber für informationswissenschaftlichen Einfluss war die Mainstream-Linguistik doch weitgehend immun. Diese Geschichte mit Modifikationen ist in einer deutschen Kurzform (Engerer 2012) und einer englischsprachigen Langform (Engerer 2016) in zwischen einem interessierten Publikum zugänglich.

Nach der Diachronie kommt die Synchronie, das wissen alle Strukturalisten. Auf die Informationswissenschaft und die Linguistik angewendet heißt dies zu fragen,

wie sich diese beiden Disziplinen jetzt zueinander verhalten, quasi systemisch. Was hält sie zusammen, was unterscheidet sie? Was kann die eine von der anderen lernen? Hierzu gibt es erste Ideen, und die sogar auf Deutsch (Engerer 2014b). Zu der Zeit, ca. 2014, war noch reichlich Fleisch in der Suppe, nämlich 4 Vokabulare (Thesauri, Terminologien, Lexika, fachsprachliche Vokabulare) und 3 Strukturprinzipien (Kontrolle, Verortung und Syntagnatisierung). Bei meiner ersten Einreichung einer englischsprachigen erweiterten Ausarbeitung im *Journal of the Association for Information Science and Technology* 2015 winkten die beiden Gutachter schnell ab, abspecken, Herr Engerer! In der gekürzten Version (Oktober 2015) fiel schon die „Verortung“ weg, und in der revidierten (und letztendlich akzeptierten) Version, wieder nach zwei anonymen Gutachten (April 2016), verschwanden Terminologien und Fachsprachen auch noch von der Bühne. Der jetzt akzeptierte und im Druck befindliche Artikel arbeitet mit zwei Vokabularen (Thesauri und natürliche Lexika) und zwei Strukturierungsprinzipien, gerade noch zu überblicken. Titel: „Control and Syntagnatization. Vocabulary Requirements in Information Retrieval Thesauri and Natural Language Lexicons“ (Engerer accepted). Das also die schlanke Version für den effektiven Informationswissenschaftler.

Ich werde hier eine „Mittelversion“ meiner Theorie anbieten, wie gehabt mit zwei Dimensionen (Kontrolle und Syntagnatisierung), aber mit der vollen Anzahl von Vokabularen (Lexika, Thesauri, Fachsprachlexika, Terminologien) wie noch 2014 (Engerer 2014b). Sprachwissenschaftler sollten ja ohne Schwindelgefühle mit vier Typen von Vokabularen gleichzeitig hantieren können.

2 Kern- und Nachbarn der Informationswissenschaft

Ziel meines synchronen Projekts ist es, ein zentrales Wissenssystem der Informationswissenschaft in ein Netzwerk mit drei anderen Disziplinen zu stellen. Hieraus ergeben sich zwei Fragen: Was ist das zentrale Wissenssystem, oder: der „Kern“ der Informationswissenschaft? Und: Was macht eine Disziplin zu einer synchronen „Nachbardisziplin“ der Informationswissenschaft (anders gesagt: Welche drei Disziplinen sollen es sein)?

Zweifelsohne ist die Informationswissenschaft ein sehr inhomogenes Feld mit vielen Unterdisziplinen (Bawden/Robinson 2012; Stock/Stock 2013). Dennoch gibt es ein einfaches Szenario, das vielen verschiedenen theoretischen und methodischen Ansätzen in dieser Disziplin explizit oder implizit zugrunde liegt: Wie wird Information in Dokumenten am besten repräsentiert und zwar so, dass sie für Benutzer findbar wird? Das ist die bekannte Information retrieval-Frage („Information retrieval“ im Folgenden abgekürzt als IR) (Chowdhury 2010). Diesem

IR-Szenario unterliegt üblicherweise eine Links-rechts-Topologie, wo der linke Pol in der Formulierung „für den Benutzer findbar“ und der rechte durch den Repräsentationsaspekt charakterisiert wird. Mit anderen Worten, dieselbe Information wird simultan von zwei Seiten angegangen: Auf der einen Seite gehen Repräsentationen als verkürzte thematische Beschreibungen vollständiger Dokumente in das Credo des Informationswissenschaftlers ein und werden in der informationswissenschaftlichen Unterdisziplin der Wissensorganisation untersucht (Hjørland 2013). Resultat dieses Verkürzungsprozesses sind Metadaten (Hider 2012) und der Prozess ihrer (professionellen) Herstellung heißt Indexierung (Lancaster 2003). Auf der anderen Seite der informationswissenschaftlichen Topologie befindet sich der Benutzer, auf den sich der IR-Aspekt im engen Sinne bezieht. Für ihn ist das Resultat ein Dokument, ‚record‘, ‚Hit‘, eine Referenz oder einfach ein Literaturhinweis, der ihm hilft, seinen Informationsbedarf zu decken. Diese beiden Seiten zu verbinden und zur Deckung zu bringen, Indexierung/Metadaten auf der einen Seite und den Suchprozess und den Output des IR-Systems auf der anderen, ist meines Erachtens das Grundproblem der Informationswissenschaft.

Nachdem ich nun eine Antwort auf die „Kernfrage“ in der Informationswissenschaft angeboten habe, ist das Problem der „Nachbardisziplin“ an der Reihe. Charakteristisch für das IR und Indexierungsprozesse im Rahmen der Dokumentrepräsentation ist, dass grundsätzlich zeichenvermittelte Realität verarbeitet und nicht direkt auf die Informationsobjekte (z. B. Textdokumente) zugegriffen wird, und bei vielen Dokumenttypen auch gar nicht zugegriffen werden kann (z. B. Bilder, Audiodateien, Gene, Menschen, ...). Indexierung und benutzerbezogener IR arbeiten stattdessen mit Vokabularen, die man auch *Indexsprachen* oder einfach *Sprachen* nennt (Svenonius 2000). Die Einheiten dieser Vokabulare heißen ‚Terme‘, ‚descriptors‘ oder, auf gut Bibliotheksdeutsch, auch „Schlagworte“. Wenn diese informationswissenschaftlichen Vokabulare intern strukturiert sind (semantische Relationen, Hierarchien, formale Restriktionen), liegt ein informationswissenschaftlicher Thesaurus vor (Engerer 2014a).

Wenn man die Vokabularbezogenheit der Informationswissenschaft im obigen grundlegenden Sinne akzeptiert, kann die Informationswissenschaft mit anderen, auch vokabularbezogenen Wissenschaften in ein synchrones Verhältnis gebracht werden. Ich arbeite hier insbesondere mit der Lexikologie (Wortsemantik), der Fachsprachenforschung und der Terminologielehre. Diese drei Disziplinen bilden meiner Meinung nach eine geeignete synchrone akademische Umgebung für die Informationswissenschaft als IR-Disziplin. Meine Argumentation ist im Wesentlichen hermeneutisch, indem ich eine plausible synchrone In-Beziehung-Setzung

der vier Disziplinen, die nun folgt, im Nachhinein als Bestätigung meiner anfänglichen Auswahl ansehe.

3 Vokabulareigenschaften

Informationswissenschaftliche Thesauri und natürlichsprachige Lexika spiegeln beide das konzeptuelle Verständnis ihrer Disziplinen in Bezug auf ihren Gegenstand wider. Beide hantieren mit Zeichen, jedoch in völlig verschiedenen Anwendungsdomänen mit spezifischen Anforderungen. Lexikalische Elemente in natürlichen Sprachen müssen speicherbar und abrufbar in mentalen (biologischen) Systemen sein, eine hinreichende Komplexität und Flexibilität für Bedeutungsvariation in der Rede aufweisen, sie sollten lernbar sein sowohl im Erst- als auch im Zweitspracherwerb (man denke hier an morphologische Transparenz) und, nicht zuletzt, ein Design vorweisen können, das den Übergang in die syntaktische Linearisierung und in kommunikative Verwendungen sichert (darüber später mehr).

Thesauruselemente leiten ihre Anforderungen und formale Kennzeichen dagegen aus dem IR-Szenario ab (Broughton 2006). Sie müssen benutzbar in Suchanfragen sowie speicherbar in digitalen Systemen sein und sich effektiv an Maschinenfeedback koppeln, so dass der Benutzer vom System die richtige Information für die interaktive Anpassung der Suche bekommt. Thesaurusterme müssen semantisch auf relevante und aktuelle Denotationen in Bezug auf die Benutzergruppe getrimmt sein und die paradigmatischen lexikalischen Relationen kodieren, die der Terminologie/Fachsprache des Benutzers entsprechen. Thesauruseinheiten sollten weiterhin in lineare Ketten eingehen, die mit einer gewissen Wahrscheinlichkeit auch in Suchabfragen vorkommen (der Aufbau von Themenbeschreibungen sollte in Indexsprache und Benutzersprache weitgehend parallel sein).

Fachsprachliche Vokabulare beziehen ihre Zeichenqualitäten aus den Postulaten der Fachsprachenforschung (vgl. die entsprechenden Artikel in Hoffmann/Kalverkämper/Wiegand 1998/1999) wie Genauigkeit (immer relativ und kontextabhängig, mehrdeutigkeitsumgehend), Explizithet (ebenso kontextabhängig, vgl. gegenseitig vorausgesetztes Wissen), Ökonomie (vs. Explizithet) oder Verständlichkeit. Terminologien strukturieren professionelles Wissen und systematisieren die Konzepte eines Wissensgebiets und deren Ausdruck. Terminologien sind kontrollierte Vokabulare mit dem Ziel der Standardisierung, Harmonisierung und der Unifikation von Terminologien. Tabelle 1 fasst das Zusammenspiel von disziplinären Gegenständen (Domänen, Problemstellungen), befassender Teiltheorien, resultierender Vokabulartypen, deren Anforderungen und nicht zuletzt

Eigenschaften der jeweiligen vokabularspezifischen Zeichen (Wörter) zusammen.

Tab. 1: Vokabulartypen nach Disziplinen geordnet

Disziplin	Domäne	Teiltheorien	Vokabulartyp	Vokabulareigenschaften	Zeichenattribute
Linguistik	Unspezifiziert	(Kognitive) Semantik, Lexikologie, ...	Lexika	Speicherbar, strukturiert, findbar in mentalen Systemen, lernbar, designed auf kommunikativen Gebrauch hin, ...	Flektierbar, Wortarten, reichhaltige semantische Relationen, deiktische Elemente, ...
Fachsprachenforschung	Spezialisierte Kommunikation, "Fachkommunikation"	Soziolinguistik, Wortschatzstruktur, ...	Fachsprachliche Lexika	Genauigkeit, Explizithet, Ökonomie, Verständlichkeit	Starke Strukturierung von domänenspezifischen Wortfeldern, ...
Informationswissenschaft	Suchkontexte, Informationsbedarf	Indexierungstheorie, Metadaten, IR, ...	Thesauri	Aktuelle Denotationen, paradigmatische Relationen, Affinität zu Benutzersynonymen, parallele Kompositivität, ...	Substantive, unflektiert, ...
Terminologielehre	Standardisierung, Harmonisierung, Unifikation, ...	Wissensstrukturierung, Wissensorganisation, ...	Terminologien	Referentielle Eindeutigkeit, ...	Verben (bezeichnen Prozesse) und nominale Einheiten, ausgeprägte semantische Relationen, ...

4 Kontrolle

Wie kann nun dieses viergliedrige System weiter ausdifferenziert werden? Meine Lösung besteht darin, zwei Dimensionen, eine nicht-kommunikative (Kontrolle) und eine sprachlich-kommunikative (Syntagmatisierung), einzuführen und zu zeigen, dass die disziplinspezifischen Vokabulareigenschaften wie in Tabelle 1 demonstriert auf diesen zwei Dimensionen systematisch variieren. Kontrolle (Svenonius 2000) ist ein gesellschaftliches Werkzeug zur Manipulation und „Pflege“ von Vokabularen und betrifft damit die soziale Sphäre mit zielorientiert

handelnden Agenten oder Institutionen. Syntagmatisierung meint die Eigenschaft von Vokabulareinheiten, in Ketten und kommunikative Äußerungen einzugehen. Es wird möglich sein, wie ich zeige, unsere vier Vokabulare entlang dieser Dimensionen zu ordnen und eine Art erklärenden Hintergrund zu liefern, wie und warum gerade sie in dieses synchrone System eingehen.

Kontrolle als intentionales, zielgerichtetes Design von Vokabularen durch menschliche Akteure ist typisch für das Design von Wissenssystemen (Batley 2005, zu Klassifikationsschemata vgl. Beghtol 2009) und IR-Systemen. Die Regeln, Prinzipien und Richtlinien der Vokabulargestaltung (thesaurus building) sind in allen Handbüchern zur Informations- und Bibliothekswissenschaft nachzulesen (z. B. Broughton 2006). Kontrolle in der Informationswissenschaft wird immer in Bezug auf bestimmte Benutzergruppen konstruiert. Man unterscheidet einmal die Gruppe der „Informationslaien“, welche die Informationssysteme nutzen, und einmal die Gruppe der Informationsprofessionellen (z. B. Bibliothekare), welche die Informationssysteme für die Informationslaien zur Verfügung stellen und pflegen. Informationslaien müssen wiederum aufgeteilt werden in unspezialisierte Benutzer (z. B. hobbybezogener oder einfach interessierter Informationsbedarf) und spezialisierte Benutzer, die ausgestattet sind mit einer Fachterminologie, wie z. B. Linguisten.

Kontrolle ist aus dieser Perspektive verhältnismäßig hoch in Bezug auf die Informationsprofessionellen, wo Vokabulare als Werkzeug zur Beschreibung von Dokumenten dienen. Informationslaien dagegen, sowohl die spezialisierten als auch die nichtspezialisierten, sind sich üblicherweise nicht der kontrollierten Beschreibungen der Informationen, die sie suchen, bewusst. Nichtspezialisierte Laien, s. o., gebrauchen in ihrer Suche Ausdrücke, welche ihrer Meinung nach das jeweilige Interessengebiet am besten beschreiben, meist mit unkontrollierten Ausdrücken aus ihrem Alltagslexikon. Spezialisierte Benutzer dagegen bringen in Recherchen typischerweise ihr Fachvokabular zur Anwendung. Gemeinsam für die Informationslaien, sowohl die spezialisierten und nicht-spezialisierten, ist, dass beide ihr Repertoire von entweder vollständig unkontrollierten Termen (z. B. der Hobbyhistoriker) oder teilweise unkontrollierten Termen (z. B. der bibliothekswingewohnte Linguist) in die Sphäre kontrollierter Thesaurussuchen transferieren. Dieser Mismatch führt zu einer Reihe von informationswissenschaftlichen Problemen und Konflikten bei der Beschreibung und dem Design von IR-Systemen, auch (oder gerade) im Zeitalter von Volltextsuchen (Hjørland 2008; Warner 2010).

Natürlichsprachige Lexika sind durch niedrige Kontrolle ausgezeichnet, welcher allerdings präskriptive und sprachpuristische Tendenzen entgegenarbeiten. Präskriptive Ansätze sind so gesehen kontrollfördernd. Terminologien beruhen auf maximaler Kontrolle, indem ein gesamtes System von Fachtermen für die fachspezifische Kommunikation (z. B. in Handbüchern für Maschinen) verbindlich eingesetzt wird. Kontrolle in Fachsprachen dagegen zeigt viele Züge von natürlichsprachigen Lexika (also niedrige Kontrolle), dennoch kommt Kontrolle durchaus auch hier Bedeutung zu und wird zu einer realistischen Möglichkeit. Fachsprachliche Kontrolle kommt in zwei Typen daher, einmal explizit und zentralisiert z. B. durch Lehrmaterialien, Betriebsanleitungen oder Kommunikationsrichtlinien, und einmal implizit, dezentralisiert. Der letztgenannte Modus, wo in metakommunikativen Ereignissen am Arbeitsplatz der „richtige“ Wortgebrauch ausgehandelt wird, ist subtiler und weniger ins Auge fallend als der explizite Typ. Es handelt sich hier um eine Art von interaktiver Kontrolle. Tabelle 2 fasst die hier erläuterten Zusammenhänge zwischen Vokabulartypen und ansteigender Kontrolle von unten nach oben noch einmal zusammen. Zu beachten ist hier der hybride Status von informationswissenschaftlichen Thesauri, wo sich 3 Kontrollgrade auf 3 Benutzergruppen hin spezifizieren lassen.

Tab. 2.: Kontrollgrad bei verschiedenen Vokabulartypen

Vokabulartyp	Kontrollgrad	Bemerkungen
Terminologien	Maximal	Vokabulare als verbindlich designte Wissenssysteme
Thesauri	Hoch – Mittel – Niedrig	<ul style="list-style-type: none"> Hoch: Informationsprofessionelle (autoritative Regulierung) Mittel: spezialisierte Benutzer, Experten (Gebrauch von spezialisiertem Vokabular in der Suchsituation) Niedrig: nichtspezialisierte „Normalbenutzer“ (Gebrauch des Alltagslexikons in Suchkontexten)
Fachsprachliche Vokabulare	Mittel	Zentralisierte Kontrolle möglich, interaktives Verhandeln der Bedeutung von Fachtermen (Metakommunikation)
Lexika	Niedrig	Unkontrolliert (vgl. aber präskriptive Tendenzen)

5 Syntagmatisierung

Ich gehe jetzt über zur sprachlich-kommunikativen Interpretation des Zusammenhangs zwischen Vokabulareigenschaften und Kontrolle. Meine These ist, dass die Bewegung von oben nach unten in Tabelle 2 (abnehmende Kontrolle) ver-

bunden ist mit einer ansteigenden Präsenz von zeichenkombinatorischen Aspekten in den jeweiligen Vokabularen. Dieses Phänomen, das ich „Syntagnatisierung“ nenne, ist auf ein Wortdesign gerichtet, das Vokabulareinheiten auf die Reihung in der syntaktischen Kette und letztendlich auf die kommunikative Verwendung dieser Einheiten im Sprechakt vorbereitet. Mit Syntagnatisierung ist damit der Übergang gemeint von einem „statischen“ System, das intrinsische, paradigmatische semantische Relationen in einem System kodiert, hinein in die „dynamische“ syntaktische Reihung in der Kette, seien es Sätze oder kleinere Konstituenten. Diese Dichotomie ist auch in Indexsprachen bekannt (Pandey 2003).

Die statisch-dynamisch-Unterscheidung ist in der Paradigmatik vs. Syntagnatik-Dichotomie weitergeführt (Happ 1985). Paradigmatische Einheiten wie Lexikoneinheiten sind verbunden durch a priori semantische Relationen, während syntagnatische Einheiten erst die Zusammensetzung neuer Bedeutungen und die Verwendung in der Kommunikation möglich machen. Auch diese Unterscheidung ist in der Theorie der Indexsprachen gut bekannt. Wie flexibel sind also Wortformen der vier Vokabulartypen, wenn es darum geht, in eine kommunizierbare syntagnatische Kette einzugehen?

Terminologien beschreiben Wissensstrukturen, die Begriffe in einem paradigmatischen Netzwerk organisieren. Diese Terme sind nicht auf die syntagnatische Reihung und Verwendung hin organisiert und als eine Sammlung lexikalischer Einheiten grammatisch gar nicht dazu ausgerüstet, Propositionen oder Prädikate zu bilden. Sie sind „Werkzeuge des Denkens“. Wenn verwendet, wird die Syntax „geliehen“ aus der Allgemeinsprache oder der Fachsprache.

Bei den Thesauri tauchen in Gestalt der sog. Zitationsordnung zum ersten Mal nicht-geliehene syntagnatische Ordnungsprinzipien auf (Broughton 2006; Batley 2005; Enger 2014a). Syntagnatische Relationen werden jedoch nicht durch Flexion oder andere dezidiert grammatische Elemente ausgedrückt, sondern durch Aureihung, Konkatenation von Indextermen ähnlich wie in isolierenden natürlichen Sprachen. Terme in Thesauri weisen keine Wortartdifferenzierung auf, die syntagnatisch nutzbar gemacht werden könnte, sie sind alle Substantive oder nominale Phrasen. Auch die Ausstattung mit semantischen Relationen ist im Vergleich zu natürlichen Sprachen (Lyons 1977) sparsam. Was die kommunikative Ebene angeht, muss der Einsatz von Indextermen in Verbindung mit dem gesamten IR-System gesehen werden. Hier führt der Benutzer mit seiner Suchanfrage einen rudimentären Frage- oder Aufforderungsakt aus, einen Sprechakttyp, der nicht an der sprachlichen Form der Suchketten abgelesen werden kann, sondern der Suchkette a priori vom Wissenssystem zugewiesen wird.

Fachsprachen haben dagegen den syntagnatischen und kommunikativen Aspekt sozusagen schon „eingebaut“, da der Kernbegriff der „fachsprachlichen Kommunikation sowie die t-esprochenen fachsprachlichen Postulate (Genauigkeit, ...) die syntaktische Reihung, den referentiellen Bezug zur Welt und die kommunikative Verwendung gewissermaßen schon voraussetzen.

Maximale Syntagnatisierung liegt in diesem System bei den natürlichen Sprachen vor. Wie bekannt, gibt es eine Reihe von grammatischen Phänomenen, die ganz ausdrücklich entweder auf die Linearisierung (Kongruenz, syntaktische Konstituentenregeln basierend auf lexikalischen Wortarten, das lexikalische Kopprinzip u. a.) oder auf den Kontext, die Situativität von Interaktion sowie auf die kommunikative Verwendung Bezug nehmen (Tempus in Bezug auf Sprechzeitpunkt, Modalität, z. B. epistemische Sprechereinstellungen, Person und Kodierung von kommunikativen Rollen, Deixis und Bezug auf die Kommunikationssituation, Diskurspartikeln und Bezug auf den Diskurs usw.). Tabelle 3 hält das Zunehmen syntagnatischer Zeichenaspekte (Syntagnatisierung) von Terminologien bis hin zu natürlichen Lexika noch einmal im Verhältnis zu abnehmender Kontrolle fest.

Tab. 3.: Syntagnatisierungsgrad bei verschiedenen Vokabulartypen

Vokabulartyp	Kontrollgrad	Syntagnatisierungsgrad	Bemerkungen
Terminologien	Maximal	Niedrig	Formal nicht für die Linearisierung vorbereitet; Wissensstrukturen, paradigmatische Netzwerke, Denkwerkzeuge
Thesauri	Hoch – Mittel – Niedrig	Mittel	Rudimentäre Anpassung an syntagnatische Kombination (Präkoordination) und kommunikativen Gebrauch (Benutzersyntax bei Suchanfragen); ansonsten fehlen z. B. Wortartklassifikation, flektive Elemente, ...
Fachsprachliche Vokabulare	Mittel	Hoch	Voll angepasst an syntaktische Kombination und Kommunikation; „Exaktheit“ z. B. setzt Situertheit in spezialisierten Kontexten voraus
Lexika	Niedrig	Hoch	Voll ausgerüstet für syntaktische Kombination und kommunikativen Gebrauch

6 Zum Schluss

Semiotische Syntagmatisierung ist an nicht-semiotische Kontrolle umgekehrt proportional gekoppelt. Warum dies? Syntaktischer Ausdruck und kommunikatives Agieren werden oft als „unkontrollierbare“ Aktivitäten betrachtet, was sich z. B. darin zeigt, dass es ein vollständiges Satzlexikon nicht geben kann, wogegen vollständige Lexika nicht nur eine theoretische Möglichkeit sind, sondern natürlich existieren. Wörter können also durchaus in einem Vokabular „kontrolliert“ werden, Sätze und Kommunikationen dagegen nicht.

Literaturverzeichnis

- Barley, Susan 2005: „*Classification in theory and practice*“. Oxford: Chandos Publishing.
- Bawden, David/Robinson, Lyn 2012: „*An introduction to information science*“. London: Facet.
- Beghel, Clare 2009: „Classification Theory“. In: *Encyclopedia of Library and Information Science*. Third Edition. Taylor & Francis, 1045–1060.
- Broughton, Vanda 2006: „*Essential thesaurus construction*“. London: Facet.
- Chowdhury, G. G. 2010: „*Introduction to modern information retrieval*“. Third Edition. London: Facet.
- Engerer, Volkmar (accepted): „Control and Syntagmatization. Vocabulary Requirements in Information Retrieval Thesauri and Natural Language Lexicons“. *Journal of the Association for Information Science and Technology*.
- Engerer, Volkmar 2016: „Exploring interdisciplinary relationships between linguistics and information retrieval from the 1960s to today“. *Journal of the Association for Information Science and Technology* Early view. 660–680.
- Engerer, Volkmar 2014a: „Indexierungstheorie für Linguisten. Zu einigen natürlichsprachlichen Zügen in künstlichen Indexsprachen“. In: Schönbeger, Manuela et al. (Hgg.): *Dialekte, Konzepte, Konstante. Ergebnisse des Arbeitsstreffens der Gesellschaft für Sprache und Sprachen, GeSuS e.V., 31. Mai – 1. Juni 2013 in Freiburg/Breisgau*. Jena: Gesis, 61–74.
- Engerer, Volkmar 2014b: „Thesauri, Terminologien, Lexika, Fachsprachen. Kontrolle, physische Verortung und das Prinzip der Syntagmatisierung von Vokabularen“. *Information, Wissenschaft & Praxis* 65. 2, 99–108.
- Engerer, Volkmar 2012: „Informationswissenschaft und Linguistik. Kurze Geschichte eines fruchtbaren interdisziplinären Verhältnisses in drei Akten“. *SDV – Sprache und Datenverarbeitung. International Journal for Language Data Processing* 36. 2, 71–91.
- Happ, Heinz 1985: „Paradigmatisch – „syntagmatisch“: zur Bestimmung und Klärung zweier Grundbegriffe der Sprachwissenschaft. Heidelberg: C. Winter.
- Hider, Philip 2012: „*Information resource description: creating and managing metadata*“. London: Facet.
- Hjørland, Birger 2013: „Theories of Knowledge Organization — Theories of Knowledge“. *Knowledge Organization* 40. 3, 169–181.
- Hjørland, Birger 2008: „What is knowledge organization (KO)?“. *Knowledge Organization* 35. 2/3, 86–101.
- Hoffmann, Lothar/Kalvertkämper, Hartwig/Wiegand, Herbert Ernst (Hgg.) 1998/1999: „*Fachsprachen: ein internationales Handbuch zur Fachsprachenforschung und Terminologiewissenschaft*“. Berlin, New York: de Gruyter.
- Lancaster, F. Wilfrid 2003: „*Indexing and abstracting in theory and practice*“. Third Edition. London: Facet.
- Lyons, John 1977: „*Semantics*“. London: Cambridge University Press.
- Pandey, R. C. 2003: „*Information retrieval system. A linguistic study*“. Delhi: Abhijeet Publications.
- Stock, Wolfgang G./Stock, Mechthild 2013: „*Handbook of information science*“. Berlin: de Gruyter Saur.
- Svenonius, Elaine 2000: „*The intellectual foundation of information organization*“. Cambridge, Mass.: MIT Press.